

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2002, 11:20:03 ; Search time 23.55 Seconds
(without alignments)
999.993 Million cell updates/sec

Title: US-09-818-066-34
Perfect score: 871
Sequence: 1 MGQHPAKSMVRRIEGGEIL.....PLYQSEPAVPVKTPLKKK 161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB ID	Description
1	871	100.0	329	12	O92935 duck hepati
2	866	99.4	364	12	Q66399 duck hepati
3	785.5	90.2	330	12	O72885 duck hepati
4	779.5	89.5	330	12	O66405 duck hepati
5	779.5	89.5	366	12	Q66404 duck hepati
6	724	83.1	329	12	O9WFB3
7	724	83.1	329	12	O9WFA9
8	724	83.1	329	12	O9WFA3
9	717	82.3	329	12	O9WFB6
10	714	82.0	329	12	O9WFA6
11	602.5	69.2	327	12	Q67852
12	98.5	11.3	418	2	Q9A4Y7
13	98	11.3	1172	4	O9Y4F2
14	96	11.0	548	4	Q9BTI7
15	96	11.0	559	4	O9BYJ9
16	95.5	11.0	315	11	O55150
17	95.5	11.0	633	10	O65655
18	95.5	11.0	1339	11	O35788
19	94.5	10.8	542	12	Q84357

20	94	10.8	1260	4	O9UGY9	O9ugy9 homo sapien
21	93.5	10.7	907	4	P78344	P78344 homo sapien
22	93.5	10.7	907	6	P79398	P79398 oryctolagus
23	92.5	10.6	269	10	O9FUR7	O9fur7 styphnolobi
24	92.5	10.6	989	11	O9JUE9	O9jie9 rattus norv
25	92	10.6	539	10	O9M2B7	O9m2b7 arabidopsis
26	92	10.6	980	12	O92809	O92809 abelson mur
27	92	10.6	981	12	O39477	O39477 abelson mur
28	91.5	10.5	299	10	O49201	O49201 gossypium h
29	91.5	10.5	1386	4	O9COA3	O9COA3 homo sapien
30	91	10.4	197	4	O9NXJ9	O9nxj9 homo sapien
31	90.5	10.4	926	3	O13305	O13305 pneumocysti
32	90	10.3	407	2	O53879	O53879 mycobacteri
33	89.5	10.3	907	11	O62448	O62448 mus musculu
34	89	10.2	283	5	O20349	O20349 caenorhabdi
35	89	10.2	593	11	O55153	O55153 rattus norv
36	89	10.2	593	11	O99MG2	O99mg2 rattus norv
37	89	10.2	616	11	O55151	O55151 rattus norv
38	89	10.2	654	11	O55152	O55152 rattus norv
39	89	10.2	654	11	O99MG1	O99mg1 rattus norv
40	88.5	10.2	493	10	O9LT85	O9lt85 arabidopsis
41	88.5	10.2	1315	10	O9SPM0	O9spm0 zea mays (m
42	88	10.1	337	11	P97268	P97268 cavia porce
43	87.5	10.0	396	4	O9H5Q4	O9h5q4 homo sapien
44	87.5	10.0	413	5	P90973	P90973 caenorhabdi
45	87.5	10.0	453	5	O02592	O02592 caenorhabdi

ALIGNMENTS

RESULT	1					
O92935						
ID	O92935	PRELIMINARY;	PRT;	329	AA.	
AC	O92935					
DT	01-NOV-1998	(TrEMBLrel. 08, Created)				
DT	01-NOV-1998	(TrEMBLrel. 08, Last sequence update)				
DT	01-JUN-2001	(TrEMBLrel. 17, Last annotation update)				
DE	PRES ANTIGEN					
OS	Duck hepatitis B virus (DHBV).					
OC	Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.					
OX	NCBI_TaxID=12639;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=ALBERTA;					
RA	Fischer K.P., Stickney J., Tipples G.A., Tyrrell D.L.J.;					
RT	"Cloning, sequencing and sequence comparison of a Canadian isolate of duck hepatitis B virus."					
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.					
DR	EMBL; AF047045; AAC06355.1;					
DR	InterPro: IPR000349; Hepadnavir_surfa9.					
DR	Pfam: PF00695; WMSA; 2.					
SO	SEQUENCE	329	AA;	36361	MW;	46E4ACAFCA4995147 CRC64;
Query Match		100.0%;	Score	871;	DB	12; Length 329;
Best Local Similarity		100.0%;	Pred. No.	6.4e-73;		
Matches	161;	Conservative	0;	Mismatches	0;	Indels
					0;	Gaps
						0;
Qy	1	MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTITWSGKFFPTLDHVDHVTMEINTLQ	60			
Db	2	MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTITWSGKFFPTLDHVDHVTMEINTLQ	61			
Qy	61	NOGAWPAGAGRRVGLSNPTPOEIPQWTPEDQKARAFRYQERPETTTTPSPSP	120			
Db	62	NOGAWPAGAGRRVGLSNPTPOEIPQWTPEDQKARAFRYQERPETTTTPSPSP	121			
Qy	121	QWKLPQGDPLLGNSLLETHPLYSQSEPAVPVKTPLKKK	161			
Db	122	QWKLPQGDPLLGNSLLETHPLYSQSEPAVPVKTPLKKK	162			
RESULT	2					

[illegible]

Query Match	89.5%;	Score	779.5;	DB	12;	Length	366;
Best Local Similarity	88.3%;	Pred. No.	2.2e-64;				
Matches	143;	Conservative	10;	Mismatches	8;	Indels	1;
Gaps	1;						
QY	1	MGQHPAKSMDVRRTEGGEILLNQLAGRMIPKGTITWGSKFPTLDHLDHVQTMEEINTLQ	60				
Db	37	MGQHPAKSMDVRRTEGGEILLNQLAGRMIPKGTITWGSKFPTLDHLDHVQTMEEINTLQ	96				
QY	61	NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTIPPSPP	120				
Db	97	QOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTIPPSPP	156				
QY	121	QWKLPQGDPLNGQSLETHPLYQS-EPAPVPVTKTPPLKKK	161				
Db	157	PKWLQPGDDPLNGQSLETHPLYQNPEAPVPVTKTPPLKKK	198				
RESULT	6						
Q9WFB3		PRELIMINARY;	PRT;	329	AA.		
ID	Q9WFB3						
AC	Q9WFB3						
DT	01-NOV-1999	(TEMBLrel. 12, Created)					
DT	01-NOV-1999	(TEMBLrel. 12, Last sequence update)					
DT	01-JUN-2001	(TEMBLrel. 17, Last annotation update)					
DE	PRES ANTIGEN.						
OS	snow goose hepatitis B virus.						
OC	Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.						
OX	NCBI_TaxID=89623;						
RN	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=SGHBV1-7;						
RA	Chang S.-F., Netter H.J., Bruns M., Schneider R., Froelich K.,						
RA	Will H.;						
RT	"Identification and functional analysis of a new avian hepatitis B						
RL	virus naturally infecting snow geese (Anser caerulescens).";						
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AF110999; AAD21997.1;						
DR	InterPro; IPR000349; Hepadnavir_surfa.						
DR	Pfam; PF00695; VMSA; 2.						
SQ	SEQUENCE	329	AA;	36602	MW;	46DAB552978B0F27	CRC64;
Query Match	83.1%;	Score	724;	DB	12;	Length	329;
Best Local Similarity	85.2%;	Pred. No.	2.6e-59;				
Matches	138;	Conservative	8;	Mismatches	14;	Indels	2;
Gaps	2;						
QY	1	MGQHPAKSMDVRRTEGGEILLNQLAGRMIPKGTITWGSKFPTLDHLDHVQTMEEINTLQ	60				
Db	1	MGQHPAKSMDARRTEGGEILLNQLAGRMIPKGTITWGSKFPTLDHLDHVQTMEEINTLQ	60				
QY	61	NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTIPPSPP	120				
Db	61	KOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTIPPSPP	119				
QY	121	QWKLPQGDPLNGQSLETHPLYQ-SEPAVPVTKTPPLKKK	161				
Db	120	PKWLQPGDDPLNGQSLETHPLYQNPEAPVPVTKTPPLKKK	161				
RESULT	7						
Q9WFA9		PRELIMINARY;	PRT;	329	AA.		
ID	Q9WFA9						
AC	Q9WFA9;						
DT	01-NOV-1999	(TEMBLrel. 12, Created)					
DT	01-NOV-1999	(TEMBLrel. 12, Last sequence update)					
DT	01-JUN-2001	(TEMBLrel. 17, Last annotation update)					
DE	PRES ANTIGEN.						
OS	snow goose hepatitis B virus.						
OC	Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.						
OX	NCBI_TaxID=89623;						
RN	[1]						

```
Q9WFB6
ID Q9WFB6 PRELIMINARY; PRT; 329 AA.
AC Q9WFB6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PRES ANTIGEN.
OS snow goose hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=89623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SGBV1-9;
RA Chang S.-F., Netter H.J., Bruns M., Schneider R., Froelich K.,
RA Will H.;
RT "Identification and functional analysis of a new avian hepatitis B
RT virus naturally infecting snow geese (Anser caerulescens).";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF111000; AAD22002.1; -
DR InterPro: IPR000349; Hepadnavir_surfaG.
DR Pfam: PF00695; VMSA; 2.
SQ SEQUENCE 329 AA; 36486 MW; 6A85C3EEEF3FA526 CRC64;

Query Match 82.3%; Score 717; DB 12; Length 329;
Best Local Similarity 84.6%; Pred. No. 1.2e-58;
Matches 137; Conservative 8; Mismatches 15; Indels 2; Gaps 2;

QY 1 MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
Db 1 MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60

QY 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTPSSPP 120
Db 61 KOGAWPEGARRVGLTNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTP-TPT 119

QY 121 QWKLPQGGDDPLLGNQSLLETHPLQ-SEPAVPVVIKTPPLKKK 161
Db 121 QWKLPQGGDDPLLGNQSLLETHPLQ-SEPAVPVVIKTPPLKKK 161

QY 120 PKWLQPGDDPLLGLTKSLLETRLQTONSEPAVPVVIKTPPLKKK 161
Db 120 PKWLQPGDDPLLGLTKSLLETRLQTONSEPAVPVVIKTPPLKKK 161

RESULT 10
Q9WFA6
ID Q9WFA6 PRELIMINARY; PRT; 329 AA.
AC Q9WFA6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE PRES ANTIGEN.
OS snow goose hepatitis B virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=89623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SGBV1-15;
RA Chang S.-F., Netter H.J., Bruns M., Schneider R., Froelich K.,
RA Will H.;
RT "Identification and functional analysis of a new avian hepatitis B
RT virus naturally infecting snow geese (Anser caerulescens).";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF110997; AAD21987.1; -
DR InterPro: IPR000349; Hepadnavir_surfaG.
DR Pfam: PF00695; VMSA; 2.
SQ SEQUENCE 329 AA; 36586 MW; 7C1928C4C0E87466 CRC64;

Query Match 82.0%; Score 714; DB 12; Length 329;
Best Local Similarity 84.0%; Pred. No. 2.2e-58;
Matches 136; Conservative 9; Mismatches 15; Indels 2; Gaps 2;

QY 1 MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
Db 1 MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
```

```
QY 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTPSSPP 120
Db 61 KOGAWPEGARRVGLTNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTP-TPT 119

QY 121 QWKLPQGGDDPLLGNQSLLETHPLQ-SEPAVPVVIKTPPLKKK 161
Db 120 PKWLQPGDDPLLGLTKSLLETRLQTONSEPAVPVVIKTPPLKKK 161

RESULT 11
Q67852
ID Q67852 PRELIMINARY; PRT; 327 AA.
AC Q67852;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SURFACE PROTEIN.
GN PRES.
OS Duck hepatitis B virus (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=12639;
RN [1]
RP SEQUENCE FROM N.A.
RA Shi H., Cullen J.M., Newbold J.E.;
RL Submitted (DEC-1992) to the EMBL/GenBank/DBJ databases.
DR EMBL: M95589; AAA45749.1; -
DR InterPro: IPR000349; Hepadnavir_surfaG.
DR Pfam: PF00695; VMSA; 2.
SQ SEQUENCE 327 AA; 36356 MW; 6875E959746DADEB CRC64;

Query Match 69.2%; Score 602.5; DB 12; Length 327;
Best Local Similarity 68.3%; Pred. No. 4.7e-48;
Matches 114; Conservative 17; Mismatches 23; Indels 13; Gaps 3;

QY 1 MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
Db 1 MGQHPAKSMVRRIEGGEILLNLAGRMIPKGTLTWSGKFPPTLDHVDHVQTMEEINTMQ 60

QY 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTPSSPP 120
Db 61 QOGAWPEGARRVGLTNPTQEIPOQWTPEDQKAREAFRRYQERPPETTTP-TPT 120

QY 121 --QWKLPQGGDDPLLGNQSLLETHPLQ----SEPAVPVVIKTPPLKKK 161
Db 121 KTWELKPGD-----PLLSTQPLYPAPPAEPDIPVVIKTPKVPKK 160

RESULT 12
Q9A4Y7
ID Q9A4Y7 PRELIMINARY; PRT; 418 AA.
AC Q9A4Y7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CONJUGAL TRANSFER PROTEIN TRBI.
GN CC2685.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Kolonay R.T., Dodson R.J., Durkin A.S., Winn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
```

RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL; AE005935; AAK24651.1; -;
DR TIGR; CC2685; -;
KW Complete proteome.
SQ SEQUENCE 418 AA; 44093 MW; 9AC05F75943954D0 CRC64;

Query Match 11.3%; Score 98.5; DB 2; Length 418;
Best Local Similarity 25.4%; Pred. No. 0.29;
Matches 44; Conservative 15; Mismatches 49; Indels 65; Gaps 8;
Qy 34 LTW-----SGKFTPLDHLVDHVQTMEEINTLON---QGAWPAGAG-----70
Db 56 LAWFALSPRALQAARAKATPVSSAVDRGTSAAEAVRQLPSDYAOGGAPPRGLPPLPGDLGRAV 115
Qy 71 ----RRVGL-----SNPTPQEIPOQ-----QMTPEEDQKAREAFRRYQER--- 107
Db 116 INHQRRDGLVVDGSDTP--MPQSTDOQAQVDAERQHQREGQARAGVYVQVTRRAE 173
Qy 108 -----PPETTTTPSSPPQWKLPQGD-----PLIGNQSLLETHPL 143
Db 174 AAVAPNLLASPEFATPPVSSPOGIGEPGDQNRKMAFMGAVPGGESIVTAHQL 226

RESULT 13
ID Q9Y4F2 PRELIMINARY; PRT; 1172 AA.
AC Q9Y4F2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE KIAA0442 PROTEIN (FRAGMENT).
GN KIAA0442.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98116655; PubMed=9455477;
RA Ishikawa K., Nagase T., Nakajima D., Seki N., Ohira M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VIII.
RT 78 new cDNA clones from brain which code for large proteins in
RT vitro.";
RL DNA Res. 4:307-313(1997).
DR EMBL; AB007902; BAA23714.1; -;
DR InterPro; IPR002965; P_Rich_extensn.
DR PRINTS; PR01217; PRICHEXTENS.
FT NON_TER 1
SQ SEQUENCE 1172 AA; 129749 MW; DCDD5B59347AFC29 CRC64;

Query Match 11.3%; Score 98; DB 4; Length 1172;
Best Local Similarity 29.6%; Pred. No. 1;
Matches 29; Conservative 13; Mismatches 42; Indels 14; Gaps 3;
Qy 73 VGLSNPT---QEIQPQWTPEDQKARE---AFRRYQERPETTTTPSPSPQWKLPQ 126
Db 206 VVLKDFQFOVAQPIQIQPQQLRAPSPDPLVQTEAPPQPLUSTQPPQGPPEAQLQP 265
Qy 127 GDDPLL-----GNQSLLETHPLYSQSEPAVPVIKTP 156
Db 266 APQVQVQRPQRPQSTQLLHQLNLPVQAHPQAQSLSQP 303
RESULT 14
ID Q9BTI7 PRELIMINARY; PRT; 548 AA.
AC Q9BTI7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE UNKNOWN (PROTEIN FOR IMAGE:3453235) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=CHORIOCARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003681; AAH03681.1; -;
FT NON_TER 1
SQ SEQUENCE 548 AA; 59695 MW; B156923EAAC01DA9 CRC64;

Query Match 11.0%; Score 96; DB 4; Length 548;
Best Local Similarity 22.4%; Pred. No. 0.68;
Matches 37; Conservative 21; Mismatches 47; Indels 60; Gaps 8;
Qy 5 PAKSMVRRRIEGGEILLNQLAGRMIPKGTLTWSGKFPTLDHVLHDHVQTMEEINTLQNOGA 64
Db 230 PAKPQPMKTKSGPV---MGGGLPP-----PPIKHMND-----IGTWDNKGK 268
Qy 65 WPAGAGRRVGLSNPTPQEIPOQWTPEDQKARE-----AFRRYQERPPPETTTIP 115
Db 269 VPKA-----PVPQAQSPQAAPQQAQPLPAQPAQQAQSPQAQSPQQAQSPQQAQSP 319
Qy 116 P-----SSPPQWKLPQGDPLIGNQSLLETHPLQ 145
Db 320 PRNRNAAFQSGGAGSDSNSPGNVQPNAP-----SVESHPVLE 358

RESULT 15
ID Q9BYJ9 PRELIMINARY; PRT; 559 AA.
AC Q9BYJ9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DJ963E22.1 (NOVEL PROTEIN SIMILAR TO NY-REN-2 ANTIGEN) (FRAGMENT).
GN DJ963E22.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd D.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL096828; CAC09391.3; -;
FT NON_TER 1
SQ SEQUENCE 559 AA; 60873 MW; DB8AA15636130E18 CRC64;

Query Match 11.0%; Score 96; DB 4; Length 559;
Best Local Similarity 22.4%; Pred. No. 0.69;
Matches 37; Conservative 21; Mismatches 47; Indels 60; Gaps 8;
Qy 5 PAKSMVRRRIEGGEILLNQLAGRMIPKGTLTWSGKFPTLDHVLHDHVQTMEEINTLQNOGA 64
Db 241 PAKPQPMKTKSGPV---MGGGLPP-----PPIKHMND-----IGTWDNKGK 279
Qy 65 WPAGAGRRVGLSNPTPQEIPOQWTPEDQKARE-----AFRRYQERPPPETTTIP 115
Db 280 VPKA-----PVPQAQSPQAAPQQAQPLPAQPAQQAQSPQAQSPQQAQSPQQAQSP 330
Qy 116 P-----SSPPQWKLPQGDPLIGNQSLLETHPLQ 145
Db 331 PRNRNAAFQSGGAGSDSNSPGNVQPNAP-----SVESHPVLE 369

Search completed: May 13, 2002, 11:23:19
Job time: 196 sec

This Page Blank (uspto)

SQ SEQUENCE 328 AA; 36230 MW; B2D771241E407456 CRC64;

```
Query Match
Best Local Similarity 100.0%; Score 871; DB 1; Length 328;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGQHPAKSMVDRRIEGEILLNOLAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
D 1 MGQHPAKSMVDRRIEGEILLNOLAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 60

QY 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQOERPPPTTIPSSPP 120
D 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQOERPPPTTIPSSPP 120

QY 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVPIKTPPLKKK 161
D 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVPIKTPPLKKK 161

RESULT 2
VMSA_HPBDC STANDARD; PRT; 365 AA.
AC P30029;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE MAJOR SURFACE ANTIGEN PRECURSOR.
GN S.
OS Duck hepatitis B virus (strain China) (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=31510;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91045091; PubMed=2235506;
RA Tong S., Mattes F., Teubner K., Blum H.E.;
RT "Complete nucleotide sequence of a Chinese duck hepatitis B virus.";
RL Nucleic Acids Res. 18:6139-6139(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M21953; AAA45746.1;
DR PIR; S12842; SAVLWE.
DR InterPro; IPR000349; Hepadnavir_surfa.
DR Pfam; PF00695; VMSA; 2.
KW Antigen.
FT PROPEP 1 198
FT CHAIN 199 365 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 365 AA; 40511 MW; 67F3A4174CB7D884 CRC64;
```

```
Query Match
Best Local Similarity 90.7%; Score 793.5; DB 1; Length 365;
Matches 147; Conservative 8; Mismatches 6; Indels 1; Gaps 1;

QY 1 MGQHPAKSMVDRRIEGEILLNOLAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
D 37 MGQHPAKSMVDRRIEGEILLNOLAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 96

QY 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQOERPPPTTIPSSPP 120
D 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQOERPPPTTIPSSPP 120

QY 97 QGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQOERPPPTTIPSSPP 156
D 97 QGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQOERPPPTTIPSSPP 156

QY 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVPIKTPPLKKK 161
D 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVPIKTPPLKKK 161

D 157 PKLQPGDDPDLGNQSLLETHPLYSQSEPAVPVPIKTPPLKKK 198
```

```
RESULT 3
VMSA_HPBDC STANDARD; PRT; 366 AA.
AC P17195;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE MAJOR SURFACE ANTIGEN PRECURSOR.
GN S.
OS Duck hepatitis B virus (white Shanghai duck isolate S31) (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=10440;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90085807; PubMed=2596031;
RA Uchida M., Esumi M., Shikata T.;
RT "Molecular cloning and sequence analysis of duck hepatitis B virus
RT genomes of a new variant isolated from Shanghai ducks.";
RL Virology 173:600-606(1989).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32991; AAA45752.1; ALT_INIT.
DR PIR; D33746; SAVLWD.
DR InterPro; IPR000349; Hepadnavir_surfa.
DR Pfam; PF00695; VMSA; 2.
KW Antigen.
FT PROPEP 1 199
FT CHAIN 200 366 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 366 AA; 40858 MW; CF60E78B7B2FCD52 CRC64;
```

```
Query Match
Best Local Similarity 89.8%; Score 782.5; DB 1; Length 366;
Matches 143; Conservative 11; Mismatches 7; Indels 1; Gaps 1;

QY 1 MGQHPAKSMVDRRIEGEILLNOLAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
D 37 MGQHPAKSMVDRRIEGEILLNOLAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 96

QY 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQOERPPPTTIPSSPP 120
D 61 NOGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQOERPPPTTIPSSPP 120

D 97 QGAWPAGARRVGLSNPTPOEIPOQWTPPEEDQKAREAFRRYQOERPPPTTIPSSPP 156
QY 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVPIKTPPLKKK 161
D 121 QWKLPQDDPDLGNQSLLETHPLYSQSEPAVPVPIKTPPLKKK 161

D 157 PKLQPGDDPDLGNQSLLETHPLYSQSEPAVPVPIKTPPLKKK 198
```

```
RESULT 4
VMSA_HPBDC STANDARD; PRT; 366 AA.
AC P17194;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 01-MAY-1991 (Rel. 18, Last annotation update)
DE MAJOR SURFACE ANTIGEN PRECURSOR.
GN S.
OS Duck hepatitis B virus (brown Shanghai duck isolate S5) (DHBV).
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=10439;
RN [1]
RP SEQUENCE FROM N.A.
```



```
RX MEDLINE-90085807; PubMed=2596031;
RA Uchida M., Esumi M., Shikata T.;
RT "Molecular cloning and sequence analysis of duck hepatitis B virus
RL genomes of a new variant isolated from Shanghai ducks.";
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32990; AAA45755.1; ALT_INIT.
DR PIR; C33746; SAVLBD.
DR InterPro; IPR000349; Hepadnavir_surfaG.
DR Pfam; PF00695; VMSA; 2.
KW Antigen.
FT PROPEP 1 199
FT CHAIN 200 366 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 170 170 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 298 298 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 366 AA; 40897 MW; 5B72879A182EFF38 CRC64;

Query Match 89.3%; Score 777.5; DB 1; Length 366;
Best Local Similarity 87.7%; Pred. No. 3.4e-56;
Matches 142; Conservative 11; Mismatches 8; Indels 1; Gaps 1;

Qy 1 MGQHPAKSMQVRRIEGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVOTMEINTLQ 60
Db 1 MGQHPAKSMQVRRIEGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVOTMEINTLQ 60
Qy 37 MGQHPAKSMQVRRIEGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVOTMEINTLQ 96
Db 37 MGQHPAKSMQVRRIEGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVOTMEINTLQ 96
Qy 61 NOGAWPACAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPSP 120
Db 61 NOGAWPACAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPSP 120
Qy 97 QGAWPACAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPSPT 156
Db 97 QGAWPACAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTPSPT 156
Qy 121 QWKLQPGDPLLGNSLLETHPLYS-EPAPVPVKTPLKKK 161
Db 121 QWKLQPGDPLLGNSLLETHPLYS-EPAPVPVKTPLKKK 161
Qy 157 PWKLQPGDPLLGNSLLETHPLYS-EPAPVPVKTPLKKK 198
Db 157 PWKLQPGDPLLGNSLLETHPLYS-EPAPVPVKTPLKKK 198

RESULT 5
VMSA_HPBHE STANDARD; PRT; 335 AA.
AC P13847; O67853;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 15-JUL-1998 (Rel. 36, Last annotation update)
DE MAJOR SURFACE ANTIGEN PRECURSOR.
GN S.
OS Heron hepatitis b virus.
OC Viruses; Retroid viruses; Hepadnaviridae; Avihepadnavirus.
OX NCBI_TaxID=28300;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-88333160; PubMed=3418788;
RA Spengel R., Kaleta E.F., Will H.;
RT "Isolation and characterization of a hepatitis B virus endemic in
RT herons.";
RL J. Virol. 62:3832-3839(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M22056; AAA45739.1;
DR EMBL; M22056; AAA45740.1; ALT_INIT.
```

```
DR PIR; B30082; SAVLHH.
DR InterPro; IPR000349; Hepadnavir_surfaG.
DR Pfam; PF00695; VMSA; 2.
KW Antigen.
FT PROPEP 1 166
FT CHAIN 167 335 MAJOR SURFACE ANTIGEN.
FT CARBOHYD 265 265 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 335 AA; 37218 MW; E90F7DA6A0623354 CRC64;

Query Match 47.2%; Score 411.5; DB 1; Length 335;
Best Local Similarity 48.5%; Pred. No. 1.5e-26;
Matches 83; Conservative 23; Mismatches 48; Indels 17; Gaps 5;

Qy 1 MGQHPAKSMQVRRIEGGELLNQLAGRMIPKGTLTWSGKFTPLDHLVHVOTMEIN 57
Db 1 MGHTQAKSTTDREVGGELLQHLAGRMIPPEFSGPITAGKFFPIQHVMDHDSVEELR 60
Qy 58 TLQNOGAWPACAGRRVGLSNPTPOEIQPQWTPEDOKARAFRRYQERPEPTTTP 116
Db 61 TLQAGGHNPGETARRRLGLDQPRTPP-PDITWTEEDKKAKEFFKQYQENRKPAAETAPP 119
Qy 117 -----SPPQWKLPQGDPLLGNSLLETHPLYS-EPAPVPVKTPLKKK 161
Db 120 ITELHAAEPQWKISP-EDPLLKAKALIPV-----KEPEVPILKVKPLTKN 164

RESULT 6
MY15_HUMAN STANDARD; PRT; 3530 AA.
AC Q9UKN7;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE MYOSIN XV (UNCONVENTIONAL MYOSIN-15).
GN MYO15.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE-20021762; PubMed=10552926;
RA Liang Y., Wang A., Belyantseva I.A., Anderson D.W., Probst F.J.,
RA Barber T.D., Miller W., Touchman J.W., Jin L., Sullivan S.L.,
RA Sellers J.R., Camper S.A., Lloyd R.V., Kachar B., Friedman T.B.,
RA Fridell R.A.;
RT "Characterization of the human and mouse unconventional myosin XV
RT genes responsible for hereditary deafness DFNB3 and shaker 2.";
RL Genomics 61:243-258(1999).
RN [2]
RP PARTIAL SEQUENCE FROM N.A., AND VARIANTS DFNB3 TYR-2111 AND PHE-2113.
RX MEDLINE-98267311; PubMed=9603736;
RA Wang A., Liang Y., Fridell R.A., Probst F.J., Wilcox E.R.,
RA Touchman J.W., Morton C.C., Morell R.J., Noben-Trauth K., Camper S.A.,
RA Friedman T.B.;
RT "Association of unconventional myosin MYO15 mutations with human
RT nonsyndromic deafness DFNB3.";
RL Science 280:1447-1451(1998).
CC -1- FUNCTION: MYOSINS ARE ACTIN-BASED MOTOR MOLECULES WITH ATPASE
CC ACTIVITY. UNCONVENTIONAL MYOSINS SERVE IN INTRACELLULAR MOVEMENTS.
CC THEIR HIGHLY DIVERGENT TAILS ARE PRESUMED TO BIND TO MEMBRANOUS
CC COMPARTMENTS, WHICH WOULD BE MOVED RELATIVE TO ACTIN FILAMENTS (BY
CC SIMILARITY). MAY PLAY A ROLE IN THE FORMATION OR MAINTENANCE OF
CC THE ACTIN-RICH STRUCTURES OF THE INNER EAR SENSORIAL HAIR CELLS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PITUITARY. ALSO EXPRESSED
CC AT LOWER LEVELS IN ADULT BRAIN, KIDNEY, LIVER, LUNG, PANCREAS,
CC PLACENTA AND SKELETAL MUSCLE. NOT EXPRESSED IN BRAIN. IN THE
CC PITUITARY, HIGHLY EXPRESSED IN ANTERIOR GLAND CELLS.
CC -1- DISEASE: DEFECTS IN MYO15 ARE THE CAUSE OF AN AUTOSOMAL RECESSIVE
CC FORM OF NONSYNDROMIC DEAFNESS (DFNB3).
CC -1- SIMILARITY: CONTAINS 1 MYOSIN-LIKE GLOBULAR HEAD DOMAIN.
```



```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01555; CAA24858.1; -.
DR PIR; A03765; QQB24.
KW Nuclear protein.
SQ SEQUENCE 938 AA; 102860 MW; 98A513AAFC0676FD CRC64;
Query Match          9.8%; Score 85.5; DB 1; Length 938;
Best Local Similarity 23.2%; Pred. No. 13;
Matches 36; Conservative 21; Mismatches 53; Indels 45; Gaps 7;
QY   25 AGRMIPKGTITWSCKF-----PTLDHVLVD-----HVOTWEINTLNQGAWP 66
      III :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB   534 AGR---RGPCVFTGDLGIESDEPASTEPEVHDQLLPAPGPDPLEIQTLPSTTSLSASP 590
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY   67 AGAGRRVGLSNPQEIQPQWTEPEEDOKAREPRRYOEERPPPTTTTTPSSPPQWKLOP 126
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB   591 SCA-----QTWPVVQSPQTPDDPTK-----QSRPPT-----AAPQWPMPL 628
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
QY   127 GDGPL--LGNSQLLETHLYQSEFAVPVPIKTPLK 159
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
DB   629 RPIPMRLMQPIFNHPGVGPTPHQTPOVEITPYK 663
      :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
RESULT 15
MEF2_MOUSE STANDARD; PRT; 514 AA.
ID MEF2_MOUSE AC Q63943; Q63944;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE MYOCYTE-SPECIFIC ENHANCER FACTOR 2D.
GN MEF2D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=94158837; PubMed=8114702;
RA Martin J.F., Milano J.M., Hustad C.M., Copeland N.G., Jenkins N.A.,
RA Olson E.N.;
RA "A Mef2 gene that generates a muscle-specific isoform via alternative
RA mRNA splicing."
RL Mol. Cell. Biol. 14:1647-1656(1994).
CC -!- FUNCTION: TRANSCRIPTION FACTOR WHICH BINDS SPECIFICALLY TO THE
CC MEF2 ELEMENT PRESENT IN THE REGULATORY REGIONS OF MANY MUSCLE-
CC SPECIFIC GENES. ACTIVATES TRANSCRIPTION VIA THIS ELEMENT. MAY BE
CC INVOLVED IN MUSCLE-SPECIFIC AND/OR GROWTH FACTOR-RELATED
CC TRANSCRIPTION.
CC -!- SUBUNIT: HETERODIMER.
CC -!- SUBCELLULAR LOCATION: NUCLEAR.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; NON-MUSCLE AND MUSCLE; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: ONE OF THE TWO SPLICE FORMS IS MUSCLE-
CC SPECIFIC.
CC -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC FACTORS. MEF2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
DR EMBL; S68893; AAB29973.1; -
DR EMBL; S68895; AAB29974.1; -
DR HSP; P11831; ISRS.
DR MGD; MGI:99533; Mef2d.
DR InterPro; IPR002100; MADS-box.
DR Pfam; PF00319; SRF-TF; 1.
DR PRINTS; PR00404; MADSDOMAIN.
DR SMART; SM00432; MADS; 1.
DR PROSITE; PS00350; MADS_BOX_1; 1.
DR PROSITE; PS00350; MADS_BOX_2; 1.
KW Transcription regulation; Nuclear protein; DNA-binding; Activator;
KW Multigene family; Alternative splicing.
FT DOMAIN 3 57 MADS.
FT DNA_BIND 58 86 MEF2-TYPE (POTENTIAL).
FT DOMAIN 252 255 POLY-PRO.
FT DOMAIN 444 449 POLY-PRO.
FT DOMAIN 3 31 ARG/LYS-RICH (BASIC).
FT DOMAIN 365 402 GLN/PRO-RICH.
FT VARSPLIC 87 132 TLKKGFGCDSPDPGDSLEOSPILLEDKYRASELDGL
FT FRYG -> ALHNDRECSPEVDEAFALTPQTEKKYKID
FT EEKKIDEEDFKMOSYRLA (IN MUSCLE
FT ISOFORM).
FT VARSPLIC 286 292 MISSING (IN MUSCLE ISOFORM).
SQ SEQUENCE 514 AA; 54992 MW; 1D25A80DB3E04F43 CRC64;

Query Match 9.88; Score 85; DB 1; Length 514;
Best Local Similarity 23.3%; Pred. No. 7.1;
Matches 47; Conservative 27; Mismatches 58; Indels 70; Gaps 11;

QY 5 PAKSMDVYRRI--EGGETILLNOLAG-----RMIPKG 32
Db 264 PSRKPDLRVITSQGKGLMHHLTGDLNNAQRLGVSTHSLTTPVSVATPSLLSQG 323
QY 33 TLTWSGKFPPTLDHVDHVTMEINTLQNGAWPAGARRVGLSNPTTPQEIPOQWTPEE 92
Db 324 -LPFS-SMPTA-VNTDYQLPSAELSLP---AFSSPAG--LALGNVTAWQPPQPPQPPQ 375
QY 93 DQKAREAFRYQERPPETTTIPSPPPQW-----KLOPGDD-PLLGNQSL 137
Db 376 PQ-----PPSQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQPPQ 422
QY 138 LETHP--LYQSEPAVPVKTTP 157
Db 423 VTTHPHISIKSEPVSPRSRSP 444
```

Search completed: May 13, 2002, 11:23:37
Job time: 194 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2002, 11:19:13 ; Search time 14.4 Seconds
(without alignments)
851.674 Million cell updates/sec

Title: US-09-818-066-34
Perfect score: 871
Sequence: 1 MQGHPAKSMVRRRIGGEIL.....PLYQSEPAVPVIKTPPLKKK 161

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	871	100.0	364	1 SAVLWD	large surface anti
2	793.5	91.1	365	1 SAVLWE	large surface anti
3	782.5	89.8	366	1 SAVLWD	large surface anti
4	777.5	89.3	366	1 SAVLWD	large surface anti
5	692	79.4	158	2 JCI095	Pre-S protein - du
6	411.5	47.2	335	1 SAVLHH	large surface anti
7	102	11.7	3530	2 A59266	unconventional myo
8	98	11.3	1172	2 T00065	hypothetical prote
9	95.5	11.0	633	2 T05005	hypothetical prote
10	92	10.6	539	2 T47405	hypothetical prote
11	92	10.6	981	1 FOMVGM	gag-abl polyprotel
12	92	10.6	1123	2 A39362	kinase-related tra
13	91.5	10.5	299	2 T09792	proline-rich prote
14	90	10.3	407	2 C70816	hypothetical prote
15	89	10.2	283	2 T16348	hypothetical prote
16	89	10.2	603	2 S15074	calpastatin - rat
17	89	10.2	654	2 T10772	calpastatin - rat
18	87	10.0	296	2 T27319	gliadin - wheat
19	86.5	9.9	309	2 T29293	hypothetical prote
20	86.5	9.9	369	2 S20500	hydroxyproline-ric
21	86.5	9.9	564	2 S37241	legumin B - fava b
22	86.5	9.9	1952	2 T48814	hypothetical prote
23	86	9.9	285	2 A41826	probable pheromone
24	86	9.9	704	1 S33704	transforming prote
25	86	9.9	1473	2 A35186	salivary agglutini
26	85.5	9.8	430	2 JC2301	hypothetical 47.8K
27	85.5	9.8	938	1 Q0BE24	nuclear antigen EB
28	85	9.8	224	2 T03238	extensin (clone Ex
29	85	9.8	280	2 T03236	extensin precursor

ALIGNMENTS

RESULT 1

SAVLWD

large surface antigen - duck hepatitis virus

N:Contains: major surface antigen; middle surface antigen

C:Species: duck hepatitis virus, DHBV

C:Date: 20-Sep-1984 #sequence_revision 20-Sep-1984 #text_change 13-Mar-1997

C:Accession: A03710; S12845

R:Mandart, E.; Kay, A.; Galibert, F.

J. Virol. 49, 782-792, 1984

A:Title: Nucleotide sequence of a cloned duck hepatitis B virus genome: comparison w

A:Reference number: A92997; MUID:84138772

A:Accession: A03710

A:Molecule type: DNA

A:Residues: 1-364 <MAN>

A:Cross-references: CB:K01834

R:Mattes, F.; Tong, S.; Teubner, K.; Blum, H.E.

Nucleic Acids Res. 18, 6140, 1990

A:Title: Complete nucleotide sequence of a German duck hepatitis B virus.

A:Reference number: S12843; MUID:91045092

A:Accession: S12845

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 36-364 <MAT>

A:Cross-references: EMBL:X12798

C:Genetics:

A:Gene: pre-S1/pre-S2/S

C:Superfamily: hepatitis B virus surface antigen

C:Keywords: glycoprotein; surface antigen

F:89-364/Product: middle surface antigen (gene pre-S2/S) #status predicted <DSA>

F:198-364/Product: major surface antigen (gene S) #status predicted <MSA>

F:32,170,296/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 871; DB 1; Length 364;
Best Local Similarity 100.0%; Pred. No. 5.2e-67;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MQGHPAKSMVRRRIGGEILLNQLAGRIPKGTLTWSGKFTLDHVLHVQVTMEINTLQ 60

Db 37 MQGHPAKSMVRRRIGGEILLNQLAGRIPKGTLTWSGKFTLDHVLHVQVTMEINTLQ 96

Qy 61 NQAWPAGAGRRVGLSNPTPOEIPQPTWPEEDQKARAFRRYQERPEPTTTPSSPP 120

Db 97 NQAWPAGAGRRVGLSNPTPOEIPQPTWPEEDQKARAFRRYQERPEPTTTPSSPP 156

Qy 121 QWKLPQGGDPLLGNSLLETHPLYSQSEPAVPVIKTPPLKKK 161

Db 157 QWKLPQGGDPLLGNSLLETHPLYSQSEPAVPVIKTPPLKKK 197

RESULT 2

SAVLWE

DB 37 MCQQPAKSM DYRRIEGGELLNLQLAGRMIPKGTVTWSGNFPTIDHLLDHVQTMEEVNTLQ 96

Best Match

Best Local Similarity 100.0%; Pred. NO. 3.5e-52;
Matches 127: Conservative 0: Mismatches 0: Indels

[illegible]

QY 21 LNOLAGRMIPKCTLT--WSGKPTFLDVLHD-----VOTMEEINTLONQAWPA 67
Db 473 VQELLEK-----TGNVTSFW-----MDHIKTHCVSVPSVEEAAATREAVYNLQ-----WPP 519
QY 68 GAGRRV-----GLSNPTQOEIPQOWTPPEEDQKAREAFRRYQERPPETTTP 115
Db 520 NGRHLLAEFVRAEEVKELEAPLP---PQHQFOAQATLSR-----PPPTALP 565
QY 116 PSSPPQWKLPQDDPLLGNSLLETHPLXOSEPAVPVTKPPL 158
Db 566 P--PP-----PLAKPPHVVERLPLPPPPPIAPEQEPPPI 597
RESULT 10
T47405
Hypothetical protein F23N14.70 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
C:Accession: T47405
R:Blocker, H.; Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat
submitted to the Protein Sequence Database, March 2000
A:Reference number: Z24465
A:Accession: T47405
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-539 <BLO>
A:Cross-references: EMBL:AL138638
A:Experimental source: cultivar Columbia; BAC clone F23N14
C:Genetics:
A:Map position: 3
A:Note: F23N14.70

Query Match 10.6%; Score 92; DB 2; Length 539;
Best Local Similarity 26.1%; Pred. No. 2.4;
Matches 37; Conservative 16; Mismatches 63; Indels 26; Gaps 7;
QY 29 IPKGTITWSGKFP--PTLDHVLHDVQTMEEINTLONQAWPAGAGRRV-----GLSNPT 79
Db 47 VPSSSATMASSFNPT--RILDHRASSHR---NRRGAFFASKRRRLVDEPIDYDLSNPA 100
QY 80 POBIPQOWTPEDQKAREAFRRYQERPPETTIPPSPQWKLPQDDPLLGNSLLE 139
Db 101 YQVLSTPLFA-----SGICSIRESLSSPPPTTS---SQPPSVSIPPSPAPPLVLSDSKD 152
QY 140 THPLXOSEPAVPVTKPPLKKK 161
Db 153 AEPAGLTNPASAP---PSPLAPK 171

RESULT 11
F0MVGW
gag-abl polyprotein - Abelson murine leukemia virus
N:Contains: amino end of core shell protein p30; core protein p15; inner coat protein p1
C:Species: Abelson murine leukemia virus
A:Note: host Mus sp. (mouse)
C:Date: 14-Nov-1983 #sequence_revision 09-Sep-1994 #text_change 11-Jun-1999
C:Accession: A03931; A00627; A93955
R:Reddy, E.P.; Smith, M.J.; Srinivasan, A.
Proc. Natl. Acad. Sci. U.S.A. 80, 3623-3627, 1983
A:Title: Nucleotide sequence of Abelson murine leukemia virus genome: structural similar
A:Reference number: A93955; MUID:83221648
A:Accession: A03931
A:Molecule type: DNA
A:Residues: 1-981 <RED>
A:Cross-references: GB:J02009; NID:g331887; PIDN:AAA46471.1; PID:g331888
A:Note: the authors translated the codon CGA for residue 186 as Glu
R:Reddy, E.P.; Smith, M.J.; Srinivasan, A.
Proc. Natl. Acad. Sci. U.S.A. 80, 7372, 1983
A:Reference number: A93980
A:Contents: annotation; erratum, residues 588-746
C:Genetics:
A:Gene: gag-abl

C:Superfamily: Abelson murine leukemia virus gag-abl polyprotein; protein kinase homo
C:Keywords: ATP; core protein; oncogene; phosphotransferase; polyprotein; transformin
F:1-131/Product: core protein p15 #status predicted <p15>
F:132-215/Product: inner coat protein p12 #status predicted <p12>
F:216-235/Region: amino end of core shell protein p30
F:248-338/Domain: SH2 homology <SH2>
F:361-621/Domain: protein kinase homology <KIN>
F:369-377/Region: protein kinase ATP-binding motif
F:392/Active site: Lys #status predicted

Query Match 10.6%; Score 92; DB 1; Length 981;
Best Local Similarity 29.0%; Pred. No. 5;
Matches 31; Conservative 12; Mismatches 38; Indels 26; Gaps 5;

QY 53 MBEINTLONQAWPAGARRVGLSNPTQOEIPQOWTPPEEDQKAREAFRRYQERPPETT 112
Db 788 MDAVNTDPTK-AGPCGE---GLRKVPVSPVKPQST-----AKPPGTP 826
QY 113 TIPPSPQWKLPQDDPLLGNO--SLLETHPLXOSEPAVPVTKPPL 157
Db 827 TSPVSTP---STAPAPSLAGDQQPSSAAFIPLISTRVSLRKTRQPP 870

RESULT 12

A39962
kinase-related transforming protein (abl) (EC 2.7.1.-) type I - mouse
N:Alternate names: protein-tyrosine kinase abl
C:Species: Mus musculus (house mouse)
C:Date: 08-Nov-1991 #sequence_revision 08-Nov-1991 #text_change 18-Jun-1999
C:Accession: A39962; A24773; S00771; A00626
R:Oppl, C.; Shore, S.K.; Reddy, E.P.
Proc. Natl. Acad. Sci. U.S.A. 84, 8200-8204, 1987
A:Title: Nucleotide sequence of testis-derived c-abl cDNAs: implications for testis-s
A:Reference number: A39962; MUID:88068561
A:Accession: A39962
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1123 <OPP>
A:Cross-references: GB:J02995; NID:g191566; PIDN:AAA88241.1; PID:g309084
A:Experimental source: clone AL6
R:Ben-Neriah, Y.; Bernards, A.; Paskind, M.; Daley, G.Q.; Baltimore, D.
Cell 44, 577-586, 1986
A:Title: Alternative 5' exons in c-abl mRNA.
A:Reference number: A90877; MUID:86133550
A:Accession: A24773
A:Molecule type: mRNA
A:Residues: 1-118 <BEN>
R:Bernards, A.; Paskind, M.; Baltimore, D.
Oncogene 2, 297-304, 1988
A:Title: Four murine c-abl mRNAs arise by usage of two transcriptional promoters and
A:Reference number: S00771; MUID:88202920
A:Accession: S00771
A:Molecule type: DNA
A:Residues: 1-26 <BER>
A:Cross-references: EMBL:X07539; NID:g49837; PIDN:CAA30411.1; PID:g49838
R:Wang, J.Y.J.; Ledley, F.; Goff, S.; Lee, R.; Groner, Y.; Baltimore, D.
Cell 36, 349-356, 1984
A:Title: The mouse c-abl locus: molecular cloning and characterization.
A:Reference number: A00626; MUID:84106840
A:Accession: A00626
A:Molecule type: DNA
A:Residues: 85-127, 'C', 129-182 <WAN>
C:Genetics:
A:Gene: abl
C:Superfamily: human protein-tyrosine kinase abl; protein kinase homology; SH2 homolo
C:Keywords: alternative splicing; ATP; autophosphorylation; nucleus; phosphoprotein;
F:68-116/Domain: SH3 homology <SH3>
F:127-217/Domain: SH2 homology <SH2>
F:240-500/Domain: protein kinase homology <KIN>
F:248-256/Region: protein kinase ATP-binding motif

Query Match 10.6%; Score 92; DB 2; Length 1123;
Best Local Similarity 29.0%; Pred. No. 5.9;
Matches 31; Conservative 12; Mismatches 38; Indels 26; Gaps 5;

QY 53 MEELINTLQNCAPAGARRVGLSNPTQEITPQQWTPTEEDQKAREAFRRYQOERPPETT 112
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 930 MDAVTDPDK-AGPGE---GLRKPVPSPVKPKQST-----AKPPGTP 968

QY 113 TTPSSPPQWKLPQGDDPLNGN--SLEETHPLYSQSEPAVPVIKTPP 157
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 969 TSPVSTP---STAPAPSLAGDQQPSAAFIPLISTRVSRKRTPP 1012

RESULT 13
T09792
proline-rich protein precursor - upland cotton
C:Species: Gossypium hirsutum (upland cotton)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 24-Nov-1999
C:Accession: T09792
R:Orford, S.J.; Timmis, J.N.
Theor. Appl. Genet. 94, 909-918, 1997
A:Title: Abundant mRNAs specific to the developing cotton fibre.
A:Reference number: Z16858
A:Accession: T09792
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-299 <ORF>
A:Cross-references: EMBL:AF044205; NID:g2829205; PID:g2829206
A:Experimental source: cultivar Siokra 1-2
C:Genetics:
A:Gene: PRP
C:Superfamily: hydroxyproline-rich glycoprotein
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-299/Product: proline-rich protein #status predicted <MAT>

Query Match 10.5%; Score 91.5; DB 2; Length 299;
Best Local Similarity 25.3%; Pred. No. 1.3;
Matches 25; Conservative 17; Mismatches 32; Indels 25; Gaps 4;

QY 78 PTQEIPOQO---WTPEEDOKAREAFRRYOERPPEITTIPSSPPQWKLPQGDDPLGN 134
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 147 PPKKEPPEPPPVYTPPKKEK-----PEPRPP--VYEPPKKPMIEPRKPPPVYTP 196

QY 135 QSLLETH-----PLYQSEPAVPVIKTPPLKKK 161
I : I : I : I : I : I : I : I : I : I : I : I : I : I : I :
Db 197 PKKEPPEKPPMVEPPKPKPMYEPKPKPPVYTPPKKEK 235

RESULT 14
C70816
hypothetical protein Rv0867c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: C70816
R:Colet, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; O'Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Ho-Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence of the pathogen H37RV
A:Reference number: A70500; MUID:98255987
A:Accession: C70816
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-407 <COL>
A:Cross-references: GB:AL022004; GB:AL123456; NID:g3261550; PIDN:CAAL17673.1; PI
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0867c
C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

```

Query Match          10.3%; Score 90; DB 2; Length 407;
Best Local Similarity 27.5%; Pred. No. 2.6;
Matches 44; Conservative 14; Mismatches 68; Indels 34; Gaps

QY 11 VRRIE-GGEILLN---QLAGRMIPKGLTWS---GKPEFTLDHVDHVQTM---EINT 58
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 47 VACESGWNWSINTGNGYJGGLOFTQS---TWAHGGGFAPSAQLASREQQIAVGERVLA 104
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 59 LONQAGWAGAGRRVGLSNPTQEIPOQWTPPEEDQKAREAFRRYQERPPETTTIPSS 118
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 105 TOGRGAWPV-CGR-GLSNATPREVLPASAAAMDAPLDAAAV-----NGEPAPLAPPAD 155
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 119 P-PQWKLPQGGDDPLLNQSLLETHPLYQSEPAVPVPIKTPP 157
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 156 PAPPVELAANDLPA-----PLGEPLPAAPADPAPP 185
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 15
T16348
hypothetical protein F42G9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16348
R:Taich, A.
submitted to the EMBL Data Library, March 1996
A:Description: The sequence of C. elegans cosmid F42G9.
A:Reference number: Z18498
A:Accession: T16348
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-283 <TAI>
A:Cross-references: EMBL:U00051; NID:gl2126305; PID:gl216306; PIDN:AAA91352.
A:Experimental source: strain Bristol N2
C:Genetics:
A:Gene: CESP:F42G9.3
A:Introns: 78/3; 179/1

Query Match          10.2%; Score 89; DB 2; Length 283;
Best Local Similarity 24.1%; Pred. No. 2;
Matches 40; Conservative 25; Mismatches 55; Indels 46; Gaps

QY 12 RRTEGGEILLNQLAGRMIPKGLT---WSGKFFTLDHVDHVQTMEEIN---TLQNGAWPA 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 51 RRVLSGE-----GRK-EKKTINFLW-----ILTLSRVDPETQYSQGSFPE 89
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 68 GAGRRVGLSNPTQEIPOQWTPPEEDQKAR-----EAFRRYQERPPETTTIP- 115
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 90 VAVEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPEPE 149
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 116 PSSSPQWKLPQGGDPLLNQSLLETHPLYQSEPAVPVPIKTPPLKKK 161
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 150 PEPEPVPEPEPEPEP-----EPEPVVKEEV-VVESPPPEQE 186
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Search completed: May 13, 2002, 11:22:32
Job time: 199 sec

```

Search completed: May 13, 2002, 11:22:32Z
Job time: 199 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2002, 11:18:23 ; Search time 12.56 seconds
(without alignments)
288.458 Million cell updates/sec

Title: US-09-818-066-34

Perfect score: 871
Sequence: 1 MGQHPAKSMVRRIEGEIL.....PLYQSEPAVPVTKTPPLKKK 161

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

- 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	871	100.0	161	2	US-08-683-262B-34
2	871	100.0	161	4	US-09-361-707-34
3	263	30.2	48	2	US-08-683-262B-68
4	263	30.2	48	4	US-09-361-707-68
5	93.5	10.7	906	2	US-08-609-230A-9
6	93.5	10.7	907	3	US-08-990-140-4
7	93.5	10.7	940	4	US-08-810-712-7
8	88.5	10.2	1194	4	US-08-538-526-1
9	88	10.1	16	2	US-08-683-262B-18
10	88	10.1	16	4	US-09-361-707-18
11	83.5	9.6	605	2	US-08-687-956A-1
12	82.5	9.5	1257	1	US-08-340-428B-49
13	82	9.4	214	1	US-08-217-327-4
14	82	9.4	1664	2	US-08-642-846-2
15	81.5	9.4	1199	1	US-08-041-538-2
16	81.5	9.4	1199	1	US-08-463-642-2
17	81.5	9.4	1199	1	US-08-455-602-2
18	81.5	9.4	1199	2	US-08-465-157-2
19	81.5	9.4	1199	5	PCT-US91-09422-2
20	81.5	9.4	1219	2	US-08-687-289A-6
21	79.5	9.1	667	4	US-09-303-064-55
22	79.5	9.1	712	1	US-08-587-889-2
23	79.5	9.1	712	2	US-08-980-060-5
24	79.5	9.1	712	4	US-09-307-185-5
25	79.5	9.1	712	5	PCT-US96-09193-2
26	79	9.1	1004	4	US-08-916-352-2
27	79	9.1	1128	1	US-08-111-939-2

28	78.5	9.0	853	3	US-09-254-325-2	Sequence 2, Appli
29	78.5	9.0	856	3	US-09-079-415-6	Sequence 6, Appli
30	78.5	9.0	856	3	US-08-750-458A-2	Sequence 2, Appli
31	78.5	9.0	1346	2	US-08-635-121-2	Sequence 2, Appli
32	78	9.0	312	3	US-08-894-017-10	Sequence 10, Appli
33	78	9.0	424	3	US-09-173-581-7	Sequence 7, Appli
34	78	9.0	424	4	US-08-894-017-23	Sequence 23, Appli
35	78	9.0	1561	3	US-08-483-533-29	Sequence 7, Appli
36	77.5	8.9	180	4	US-08-483-533-41	Sequence 29, Appli
37	77.5	8.9	355	4	US-08-483-533-41	Sequence 41, Appli
38	77.5	8.9	355	5	PCT-US91-06532-3	Sequence 3, Appli
39	77.5	8.9	396	3	US-09-082-310-2	Sequence 2, Appli
40	77.5	8.9	550	3	US-08-659-188-18	Sequence 18, Appli
41	77.5	8.9	550	3	US-08-655-227-18	Sequence 18, Appli
42	77.5	8.9	550	3	US-08-655-241-18	Sequence 18, Appli
43	77.5	8.9	746	3	US-09-144-759-18	Sequence 18, Appli
44	77.5	8.9	764	3	US-09-144-759-20	Sequence 20, Appli
45	77	8.8	276	3	US-08-633-993A-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-08-683-262B-34
; Sequence 34, Application US/08683262B
; Patent No. 5929220
; GENERAL INFORMATION:
; APPLICANT: Shuping Tong et al.
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,262B
; FILING DATE: 18-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 31,819
; REFERENCE/DOCKET NUMBER: 00786/287002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-683-262B-34

Query Match 100.0%; Score 871; DB 2; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.3e-85;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGQHPAKSMVRRIEGEILLNOLAGRMIPKGLTWSGKPTLDHVLHDVOTMEEINTLQ 60
Db 1 MGQHPAKSMVRRIEGEILLNOLAGRMIPKGLTWSGKPTLDHVLHDVOTMEEINTLQ 60
Qy 61 NOGAWPAGARRVGLSNPTPOEIPQWTPEDQKAREAFRRYQEEPRPETTTTPSPSP 120

```
Db 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQEEPPETTTIPSSPP 120
;
QY 121 QWKLPQDDPDLGNQSLLETHPLYQSEPAVPVTKTPPLKKK 161
;
Db 121 QWKLPQDDPDLGNQSLLETHPLYQSEPAVPVTKTPPLKKK 161
;
RESULT 2
US-09-361-707-34
; Sequence 34, Application US/09361707
; Patent No. 6258937
; GENERAL INFORMATION:
; APPLICANT: Tong, Shuping
; Li, Jisu
; Wands, Jack R.
; TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,707
; FILING DATE: 27-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/683,262
; FILING DATE: 18-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 00786/287003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-8906
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 161 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 34:
US-09-361-707-34
;
Query Match 100.0%; Score 871; DB 4; Length 161;
Best Local Similarity 100.0%; Pred. No. 2.3e-85;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 1 MGQHPAKSMVRRIGEGEILLNLAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
;
Db 1 MGQHPAKSMVRRIGEGEILLNLAGRMIPKGLTWSGKFPPTLDHVDHVQTMEEINTLQ 60
;
QY 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQEEPPETTTIPSSPP 120
;
Db 61 NOGAWPAGARRVGLSNPTQEIPOQWTPEDQKAREAFRRYQEEPPETTTIPSSPP 120
;
QY 121 QWKLPQDDPDLGNQSLLETHPLYQSEPAVPVTKTPPLKKK 161
;
Db 121 QWKLPQDDPDLGNQSLLETHPLYQSEPAVPVTKTPPLKKK 161
;
RESULT 3
US-08-683-262B-68
```

```
; Sequence 68, Application US/08683262B
; Patent No. 5929220
; GENERAL INFORMATION:
; APPLICANT: Shuping Tong et al.
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,262B
; FILING DATE: 18-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 31,819
; REFERENCE/DOCKET NUMBER: 00786/287002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 68:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 48 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; US-08-683-262B-68
;
Query Match 30.2%; Score 263; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.4e-21;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
QY 69 AGRVGLSNPTQEIPOQWTPEDQKAREAFRRYQEEPPETTTIPP 116
;
Db 1 AGRVGLSNPTQEIPOQWTPEDQKAREAFRRYQEEPPETTTIPP 48
;
RESULT 4
US-09-361-707-68
; Sequence 68, Application US/09361707
; Patent No. 6258937
; GENERAL INFORMATION:
; APPLICANT: Tong, Shuping
; Li, Jisu
; Wands, Jack R.
; TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,707
; FILING DATE: 27-Jul-1999
; PRIOR APPLICATION DATA:
```


APPLICATION NUMBER: 08/683,262
FILING DATE: 18-JUL-1996
ATTORNEY/AGENT INFORMATION:
NAME: Creason, Gary L.
REGISTRATION NUMBER: 34,310
REFERENCE/DOCKET NUMBER: 00786/287003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
INFORMATION FOR SEQ ID NO: 68:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 68:
US-09-361-707-68

Query Match 30.2%; Score 263; DB 4; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.4e-21;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 69 AGRVCLSNPTQEIPOQWTFEEDQKAREAFRRYQEEPPETTTIPP 116
Db 1 AGRVCLSNPTQEIPOQWTFEEDQKAREAFRRYQEEPPETTTIPP 48

RESULT 5
US-08-609-230A-9
; Sequence 9, Application US/08609230A
; Patent No. 5866333
; GENERAL INFORMATION:
; APPLICANT: Innerarity, Thomas L.
; APPLICANT: Qian, Xiaobing
; APPLICANT: Yamanaka, Shinya
; TITLE OF INVENTION: Screening Methods to Detect mRNA Targets
; TITLE OF INVENTION: of Editing Enzymes
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/609,230A
; FILING DATE: 01-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitts, Renee A.
; REGISTRATION NUMBER: 35,136
; REFERENCE/DOCKET NUMBER: 023070-06810005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-326-2400
; TELEFAX: 650-326-2422
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 906 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-609-230A-9

Query Match 10.7%; Score 93.5; DB 2; Length 906;
Best Local Similarity 25.3%; Pred. No. 0.14;

Matches 44; Conservative 30; Mismatches 59; Indels 41; Gaps 10;

QY 1 MGQHPAKSMQVRRRIEGEIL-----LNQLAGRMIPKGTLTWSGKFPPTLDHVDHVTME 54
Db 397 MGRH--RSNQLFNGHGHIMPTQSQFGEMGKFKMSQGLS-----OLYHNQSG 444
QY 55 EINTLQNGAMPAGARRVGLSNPTQEIPOQWTFEE--DQKAREAFRRYQEEPPETTTIPP 110
Db 445 LLSQLQSGS-----KDMPPFSKKGQLNADEISLRPAQSFLMKNKQVPLQPOQ 492
QY 111 TTTIPPSS--PQWKLOP-GDDPLGNQSLLETH-PLYQSEPAVPVVIKTPPLKKK 161
Db 493 ITMIPPSAQPPRTQTPPLGQTPQLG----LKTNPPLIQEKPAKTSKKPPPSKEE 542

RESULT 6
US-08-990-140-4
; Sequence 4, Application US/08990140A
; Patent No. 6093795
; GENERAL INFORMATION:
; APPLICANT: Olsen, Henrik S.
; APPLICANT: Ruben, Steven M.
; APPLICANT: Sonenberg, Nahum
; APPLICANT: Methot, Nathalie
; APPLICANT: Rom, Eran
; TITLE OF INVENTION: Human Prt1-like Subunit Protein (hprt1) and Human
; FILE REFERENCE: 1488.0700001
; CURRENT APPLICATION NUMBER: US/08/990,140A
; CURRENT FILING DATE: 1997-12-12
; EARLIER APPLICATION NUMBER: US 60/033,151
; EARLIER FILING DATE: 1996-12-13
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-990-140-4

Query Match 10.7%; Score 93.5; DB 3; Length 907;
Best Local Similarity 25.3%; Pred. No. 0.14;
Matches 44; Conservative 30; Mismatches 59; Indels 41; Gaps 10;

QY 1 MGQHPAKSMQVRRRIEGEIL-----LNQLAGRMIPKGTLTWSGKFPPTLDHVDHVTME 54
Db 398 MGRH--RSNQLFNGHGHIMPTQSQFGEMGKFKMSQGLS-----OLYHNQSG 445
QY 55 EINTLQNGAMPAGARRVGLSNPTQEIPOQWTFEE--DQKAREAFRRYQEEPPETTTIPP 110
Db 446 LLSQLQSGS-----KDMPPFSKKGQLNADEISLRPAQSFLMKNKQVPLQPOQ 493
QY 111 TTTIPPSS--PQWKLOP-GDDPLGNQSLLETH-PLYQSEPAVPVVIKTPPLKKK 161
Db 494 ITMIPPSAQPPRTQTPPLGQTPQLG----LKTNPPLIQEKPAKTSKKPPPSKEE 543

RESULT 7
US-08-810-712-7
; Sequence 7, Application US/08810712G
; Patent No. 6160106
; GENERAL INFORMATION:
; APPLICANT: Yeda Research and Development Co. LTD
; TITLE OF INVENTION: Tumor Suppressor Genes, Proteins Encoded Thereby and
; TITLE OF INVENTION: Use of said Genes and Proteins
; FILE REFERENCE: sequence list
; CURRENT APPLICATION NUMBER: US/08/810,712G
; CURRENT FILING DATE: 1997-03-03
; EARLIER APPLICATION NUMBER: PCT/US94/11598
; EARLIER FILING DATE: 1994-10-12
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1

RESULT 12
US-08-340-428B-49
Sequence 49, Application US/08340428B
Patent No. 5648465
GENERAL INFORMATION:
APPLICANT: MARGOLIS, Richard U.
APPLICANT: RAUCH, Uwe
APPLICANT: MARGOLIS, Renee K.
TITLE OF INVENTION: CLONING, EXPRESSION AND USES FOR A
PROTEOGLYCAN
TITLE OF INVENTION: NEUROCAN AS A CHONDROITIN SULFATE
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,428B
FILING DATE: 14 No. 5648465ember 1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/922,911
FILING DATE: 03 August 1992
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: Margolis-la
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 49:
SEQUENCE CHARACTERISTICS:

US-08-041-538-2
 ; Sequence 2, Application US/08041538
 ; Patent No. 5385831
 ; GENERAL INFORMATION:
 ; APPLICANT: Mulvihill, Eileen R
 ; APPLICANT: Hagen, Frederick S
 ; APPLICANT: Houamed, Khaled M
 ; APPLICANT: Almers, Wolfhard
 ; TITLE OF INVENTION: G PROTEIN COUPLED GLUTAMATE
 ; TITLE OF INVENTION: RECEPTORS
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend
 ; STREET: Steuart Street Tower, One Market Plaza
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.24
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/041,538
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/648,481
 ; FILING DATE:
 ; APPLICATION NUMBER: US 07/626,806
 ; FILING DATE: 12-DEC-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Parmelee, Steven W
 ; REGISTRATION NUMBER: 31,990
 ; REFERENCE/DOCKET NUMBER: 13952-6-1
 ; TELEPHONE: 206-467-9600
 ; TELEFAX: 206-623-6793
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1199 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-041-538-2

Query Match 9.4%; Score 81.5; DB 1; Length 1199;
 Best Local Similarity 22.9%; Pred. No. 3.9;
 Matches 36; Conservative 19; Mismatches 65; Indels 37; Gaps 5;
 Qy 16 GGEILLNQLAGRMIPKGTLTWSGRFPTLDHVDHVTMEETLQNGA-----WPAG 68
 Db 890 GKSWSWSEPGGQAPKQGHVW-----QRLSVHVKTNE---TACNQTAIVIKPLTKSYQG 939
 Qy 69 AGRRVGLSNPTQEIPOQWTPEDQKARFRYOEERPPETTTIPSPSPQWKLOPG- 127
 Db 940 SGKSLTFS-----DASTKTLNVEEDNTPSAHFSPSPSMVVRHRCGP 983
 Qy 128 ---DDPLGNQSLLETHLYOSEPAVPIKTPPLKKK 161
 Db 984 PVATTPLPLHETAEETPLFLADSVIPKGLPPLPQQ 1020

Search completed: May 13, 2002, 11:20:18
 Job time: 115 sec

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 2002, 11:16:28 ; Search time 23.45 Seconds
(without alignments)
508.563 Million cell updates/sec

Title: US-09-818-066-34
Perfect score: 871
Sequence: 1 MGQHPAKSMDVRRIEGEL.....PLYQSEPAVPVIRKPLKKK 161

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %	Score	Match	Length	ID	Description
1	871	100.0	161	18	AAW11854	Duck hepatitis B v
2	95.5	11.0	424	21	AAG31344	Arabidopsis thalia
3	95.5	11.0	559	21	AAG31343	Arabidopsis thalia
4	95.5	11.0	633	21	AAG31342	Arabidopsis thalia
5	93.5	10.7	906	18	AAW25681	Transgenic mouse N
6	93.5	10.7	907	19	AAW49032	Human eIF4G-like p
7	93.5	10.7	940	19	AAW71370	Death associated p
8	93.5	10.7	940	21	AAW58970	Breast and ovarian
9	90	10.3	407	20	AAW95400	M. tuberculosis RP
10	90	10.3	407	22	AAW52470	Mycobacterium tube
11	89	10.2	1319	22	AAW40120	Human polypeptide

12	88.5	10.2	1194	14	AAW42199	HSmGluR1. Homo sa
13	87.5	10.0	396	22	AAW38840	Human polypeptide
14	87.5	10.0	407	22	AAW40626	Human polypeptide
15	86.5	9.9	528	21	AAW23501	Human negative elo
16	86.5	9.9	528	22	AAW47148	CDIFF-22, Incyte I
17	86	9.9	407	22	AAW91679	C glutamicum prote
18	85.5	9.8	336	19	AAW79094	Human secreted pro
19	85.5	9.8	337	20	AAW07754	Human secreted pro
20	85	9.8	280	14	AAW42055	Insecticidal/Fung1
21	85	9.8	377	22	AAW03643	Human extracellular
22	84.5	9.7	520	16	AAW78185	Protein sequence o
23	84	9.6	417	20	AAW97698	Staphylococcus aur
24	84	9.6	417	20	AAW97699	Staphylococcus aur
25	84	9.6	417	20	AAW97700	Staphylococcus aur
26	83.5	9.6	605	16	AAW79625	Endocarditis speci
27	83	9.5	442	22	AAW92045	C glutamicum prote
28	83	9.5	442	22	AAW79834	Corynebacterium gl
29	82.5	9.5	458	21	AAW70512	Caenorhabditis ele
30	82.5	9.5	653	13	AAW27150	PspA fragment. St
31	82.5	9.5	1257	15	AAW46627	Neurocan core prot
32	82	9.4	214	17	AAW86913	Cotton fiber-speci
33	82	9.4	274	11	AAW04119	ORF2 of Enod2b gen
34	82	9.4	309	11	AAW04116	ORF1 of Enod2a gen
35	82	9.4	1664	20	AAW99462	C.albicans alpha-1
36	81.5	9.4	539	22	AAW92826	Human protein sequ
37	81.5	9.4	1219	18	AAW25763	Amino acid sequenc
38	81.5	9.4	2665	22	AAW14533	Peptide #967 encod
39	81.5	9.4	2665	22	AAW26950	Peptide #987 encod
40	81.5	9.4	2665	22	AAW02259	Peptide #941 encod
41	81.5	9.4	3266	21	AAW42491	Human ORFX ORF2255
42	80.5	9.2	340	22	AAW38727	Human polypeptide
43	80.5	9.2	340	22	AAW64370	Amino acid sequenc
44	80.5	9.2	465	22	AAW40513	Human polypeptide
45	80.5	9.2	484	8	AAW71081	Sequence encoded b

ALIGNMENTS

RESULT	1	
AAW11854		
ID	AAW11854 standard; Protein; 161 AA.	
XX		
AC	AAW11854;	
XX		
DT	10-MAY-1997 (first entry)	
XX		
DE	Duck hepatitis B virus pre-S domain.	
XX		
KW	Hepadnavirus receptor; p120; p170; vaccine; pre-S domain.	
XX		
OS	Duck hepatitis B virus.	
XX		
EH	Key	Location/Qualifiers
FT	Binding-site	87..102
FT		/label= p170_binding_site
FT		/note= "the p170 binding site has been mapped
FT		to a major neutralising epitope of the
FT		pre-S domain (aa87-102), within which
FT		the Lys-95 and Arg-97 residues required
FT		for virion-receptor interaction"
FT	Peptide	1..102
FT		/note=
FT		"claimed pre-S fragment (Claim 20)
FT		capable of binding hepadnavirus receptor"
FT	Peptide	1..104
FT		/note=
FT		"claimed pre-S fragment (Claim 20)
FT		capable of binding hepadnavirus receptor"
FT	Peptide	1..126
FT		/note=
FT		"claimed pre-S fragment (Claim 20)
FT		capable of binding hepadnavirus receptor"
FT	Peptide	1..138
FT		/note=
FT		"claimed pre-S fragment (Claim 20)
FT		capable of binding hepadnavirus receptor"

FT Peptide 25..102 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 25..104 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 25..126 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 25..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 42..102 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 59..104 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 59..126 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 59..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 71..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 71..126 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 80..104 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 80..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 87..161 /note= "claimed pre-S fragment (Claim 20)
FT capable of binding hepadnavirus receptor"
FT Peptide 87..104 /note= "claimed pre-S fragment (Claim 21)
FT capable of binding hepadnavirus receptor"
XX WO9704000-A1.
XX 06-FEB-1997.
XX 22-JUL-1996; 96WO-US12098.
XX 21-JUL-1995; 95US-0001371.
XX (GEO) GEN HOSPITAL CORP.
XX Li J, Tong S, Wands JR;
XX WPI; 1997-132572/12.
XX N-PSDB; AAT59583.
XX p170, a new avian hepadnavirus receptor - binds to pre-S domain of
PT duck hepatitis B virus, attenuated p170 may be used to immunise
PT animals against hepadnaviral infection
XX Claim 19; Page 87-88; 175pp; English.
XX The pre-S domain (AAW11854) of the duck hepatitis B virus (DHBV)
CC large envelope protein binds to the cellular p170 receptor (see
CC also AAW11851-52) at a major neutralising epitope, within which are 2
CC basic amino acids (lys-95, Arg-97) required for virion-receptor
CC interaction. Peptide sequences encompassing lys-95 and Arg-97
CC can act as p170 ligands. Such peptides can be used to reduce
CC the level of hepadnaviral infection in an animal. The pre-S
CC domain (esp. attenuated), and nucleic acids encoding it, can also
CC be used as vaccines to immunise animals against hepadnavirus

CC infection.
XX Sequence 161 AA;
SQ
Query Match 100.0%; Score 871; DB 18; Length 161;
Best local Similarity 100.0%; Pred. No. 6e-76;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGQHPAKSMVRRIGEGEILLNOLAGRMIPKGTLTWSGKFTLDHVDHVTMEINTLQ 60
|||||
DB 1 mgqhpaksmvrrigeggeillnqlagrmipkgtltwsghfptldhvdhvtmeeintliq 60
|||||
QY 61 MGCAWPAGAGRRVGLSNPTPEIQPQWTPEDQKAREAFRRYQBERPPTTTPSSPP 120
|||||
DB 61 nggawpagagrrvglnsnptpgeipqpwtpedqkareafrryqgeerpttippssp 120
|||||
QY 121 QNKLQPCDDPLGNQSLLETHPLQSEPAVPVTKPPLKKK 161
|||||
DB 121 qwklpgddplgnqsllethplqsepavpvtkpplkkk 161
RESULT 2
AAG31344
ID AAG31344 standard; Protein; 424 AA.
XX AAG31344;
XX
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37626.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 20-MAY-1999; 99US-0146386.
PR 21-MAY-1999; 99US-0146388.
PR 24-MAY-1999; 99US-0146389.
PR 25-MAY-1999; 99US-0147038.
PR 27-MAY-1999; 99US-0147204.
PR 28-MAY-1999; 99US-0147302.
PR 01-JUN-1999; 99US-0147192.
PR 03-JUN-1999; 99US-0147260.
PR 04-JUN-1999; 99US-0147303.
PR 07-JUN-1999; 99US-0147416.
PR 08-JUN-1999; 99US-0147433.
PR 10-JUN-1999; 99US-0147935.
PR 10-JUN-1999; 99US-0148171.
PR 14-JUN-1999; 99US-0148319.
PR 16-JUN-1999; 99US-0148341.
PR 17-JUN-1999; 99US-0148565.
PR 18-JUN-1999; 99US-0148684.
PR 18-JUN-1999; 99US-0149368.
PR 18-JUN-1999; 99US-0149175.
PR 18-JUN-1999; 99US-0149426.
PR 18-JUN-1999; 99US-0149722.
PR 18-JUN-1999; 99US-0149723.
PR 18-JUN-1999; 99US-0149929.
PR 18-JUN-1999; 99US-0149902.
PR 18-JUN-1999; 99US-0149930.
PR 18-JUN-1999; 99US-0150566.
PR 18-JUN-1999; 99US-0150884.
PR 18-JUN-1999; 99US-0151065.
PR 18-JUN-1999; 99US-0151066.
PR 18-JUN-1999; 99US-0151080.
PR 18-JUN-1999; 99US-0151303.
PR 18-JUN-1999; 99US-0151438.
PR 18-JUN-1999; 99US-0151930.
PR 18-JUN-1999; 99US-0152363.
PR 18-JUN-1999; 99US-0153070.
PR 18-JUN-1999; 99US-0153758.
PR 18-JUN-1999; 99US-0154018.
PR 18-JUN-1999; 99US-0154039.
PR 18-JUN-1999; 99US-0154779.
PR 18-JUN-1999; 99US-0155139.
PR 18-JUN-1999; 99US-0155486.
PR 18-JUN-1999; 99US-0155659.
PR 18-JUN-1999; 99US-0156458.
PR 18-JUN-1999; 99US-0156596.
PR 18-JUN-1999; 99US-0157117.
PR 18-JUN-1999; 99US-0157753.
PR 18-JUN-1999; 99US-0157865.
PR 18-JUN-1999; 99US-0158029.
PR 18-JUN-1999; 99US-0158232.
PR 18-JUN-1999; 99US-0158369.
PR 18-JUN-1999; 99US-0159293.
PR 18-JUN-1999; 99US-0159294.
PR 18-JUN-1999; 99US-0159295.
PR 18-JUN-1999; 99US-0159329.
PR 18-JUN-1999; 99US-0159330.
PR 18-JUN-1999; 99US-0159331.
PR 18-JUN-1999; 99US-0159637.
PR 18-JUN-1999; 99US-0159638.
PR 18-JUN-1999; 99US-0159584.
PR 18-JUN-1999; 99US-0160741.
PR 18-JUN-1999; 99US-0160767.
PR 18-JUN-1999; 99US-0160768.
PR 18-JUN-1999; 99US-0160770.
PR 18-JUN-1999; 99US-0160814.
PR 18-JUN-1999; 99US-0160815.
PR 18-JUN-1999; 99US-0160980.
PR 18-JUN-1999; 99US-0160981.
PR 18-JUN-1999; 99US-0160989.
PR 18-JUN-1999; 99US-0161404.
PR 18-JUN-1999; 99US-0161405.
PR 18-JUN-1999; 99US-0161406.
PR 18-JUN-1999; 99US-0161359.
PR 18-JUN-1999; 99US-0161360.


```

PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 11.0%; Score 95.5; DB 21; Length 559;
Best Local Similarity 24.5%; Pred. No. 0.68;
Matches 40; Conservative 15; Mismatches 45; Indels 63; Gaps 10;

QY 21 LNQLAGRMIPKGTLLT--WSGKPTFLDHVLDH-----VQTMEINTLQNGAWPA 67
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 399 vqelllgk---tgnvtsfw-----mdhktchcyvspsveaaatreaavynlq----wpp 445
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 68 GAGRRV-----GLSNPTQEI PQOWTPEEDQKAREFRRYOERPPETTTIP 115
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 446 nggrhlaiefvraeevkekaleaplp---pqphqpaqatlsr-----ppptalp 491
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 116 PSSPPQWKLPQGDPLLNQSLLETHPLYQSEPAVPVVIKTPPL 158
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 492 p--pp-----plakpvhvverlpplpppplapeeqeppi 523

RESULT 4
AAG31342
ID AAG31342 standard; Protein; 633 AA.
XX
AC AAG31342;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 37624.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 19-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.

```

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

```
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161923.
PR 29-OCT-1999; 99US-0162142.

Query Match 11.0%; Score 95.5; DB 21; Length 633;
Best Local Similarity 24.5%; Pred. No. 0.79;
Matches 40; Conservative 15; Mismatches 45; Indels 63; Gaps 10;

Qy 21 LNQLAGRMIPKGTLT--WSGRFPTLDHVLDH-----VQTMETINTLQOGAWPA 67
Db 473 vqellgk---tgnvtstf-----mdhikthcyvspsveaaatreaavnlq-----wpp 519

Qy 68 GAGRRV-----GLSNPTQEIPOQWTPPEEDQKAREAFRRYQERPPETTTIP 115
Db 520 nggrhllaeafvraeevkeleaplp---pqphqpqaqlsr-----ppptaip 565

Qy 116 PSSPPQMKLQPDGDDPLLGNSQLLETHPLYOSEPAVPVVIKTPPL 158
Db 566 p--pp-----plakpphvvrrlpplppppiapeegeppi 597

RESULT 5
AAW25681
ID AAW25681 standard; Protein; 906 AA.
AC AAW25681;
DT 31-MAR-1998 (first entry)
DE Transgenic mouse NTA1 polypeptide.
KW RNA editing; mooring primer; screening; detection; editing enzyme;
KW obesity; APOBEC-1; therapeutic; NTA1; transgenic.
OS Synthetic.
OS Homo sapiens.
OS Mus sp.
XX
XX
XX Key Location/Qualifiers
XX Protein 1..906
XX /label= NTA1
XX /note= "partial coding sequence"
XX
XX WO9737032-A1.
XX
XX 04-SEP-1997.
XX
XX 28-FEB-1997; 97WO-US03189.
XX
XX 01-MAR-1996; 96US-0609230.
XX (REGC ) UNIV CALIFORNIA.
XX
XX Innerarity TL, Qian X, Yamanaka S;
XX
XX WPI; 1997-448694/41.
XX N-PSDB; AAT86087.
XX
XX Detection of mRNA targets for editing enzymes - by amplifying RNA
XX from tissues by RT-PCR using mooring primers to identify genes
XX responsible for non-wild type phenotype(s)
XX
XX Example 4; Figure 5; 36pp; English.
XX
XX This sequence represents a novel target of the apoB mRNA-editing
XX enzyme catalytic polypeptide #1, APOBEC-1, (i.e NTA1) which is used
XX as an example of a novel method of detecting mRNA candidates for editing
XX in a tissue. The method involves the generation of cDNA from RNA in the
XX tissue using one or more mooring primers and reverse transcriptase (RT)
```

```
CC and is amplified by PCR using one or more mooring primers and a 5'
CC arbitrary primer. The method has been used for identification of an
CC obesity gene e.g. APOBEC-1 or an oncogene in an animal and also for the
CC identification of non-wild type phenotypes. The identification of such
CC mRNA's provides insight into the role of the genes encoding them, in
CC growth, differentiation, or lipid accumulation, and also into the
CC possible role of an RNA editing enzyme or polypeptide in causing obesity
CC or as an oncogene. The identified polypeptides can be used as therapeutic
CC reagents in situations where the edited version of the polypeptide
CC results in a pathological state. Additionally, such polypeptides or the
CC edited versions can be used in the preparation of antibodies for
CC therapeutic use.
XX
XX SQ Sequence 906 AA;

Query Match 10.7%; Score 93.5; DB 18; Length 906;
Best Local Similarity 25.3%; Pred. No. 1.9;
Matches 44; Conservative 30; Mismatches 59; Indels 41; Gaps 10;

Qy 1 MGQHPAKSMDEVRIEGGEIL-----LNQLAGRMIPKGTLTWSGRFPTLDHVLHVQTM 54
Db 397 mgrh--rsnqlfngghimpptqsgfgemgkfkmsqgls-----qlyhnqsgg 444

Qy 55 EINTLQOGAWPAGRRVGLSNPTQEIPOQWTPPEE-DOKAREAFRRYQERP---PE 110
Db 445 llsqllqggs-----kdmpprfsskkgqlnadeislrpaqsfImnknqvpklqpg 492

Qy 111 TTTTIPSS-PPQWKLOP-GDDPLLGNSQLLETH-PLYOSEPAVPVVIKTPPLKKK 161
Db 493 itmppsaaqprrtqpgtqpgig----lknppllqekpaktsskpppskee 542

RESULT 6
AAW49032
ID AAW49032 standard; Protein; 907 AA.
XX
XX AC AAW49032;
XX
XX 07-OCT-1998 (first entry)
XX
XX Human eIF4G-like protein (p97).
XX
XX Human Prt1-like subunit protein; hPrt1; transcription; apoptosis;
XX p97; antagonist; agonist; human eIF4G-like protein.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 788..802
XX /note= "INT peptide used to raise anti-INT
XX antibodies"
XX
XX WO9825957-A2.
XX
XX 18-JUN-1998.
XX
XX 12-DEC-1997; 97WO-US22664.
XX
XX 13-DEC-1996; 96US-0033151.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX (UYMC-) UNIV MCGILL.
XX
XX Inataka H, Methot N, Olsen HS, Rom E, Ruben SM;
XX Sonenberg N;
XX
XX WPI; 1998-348451/30.
XX N-PSDB; AAV32791.
XX
XX Polynucleotides encoding human hPrt1 and p97 - useful for treating
XX disease states associated with apoptosis
XX
```


XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antiinflammatory; antitumor; antitumor; anticonvulsant;
 CC antibacterial; antifungal; antiparasitic and cardiant activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemias; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.

XX Sequence 940 AA;

Query Match 10.7%; Score 93.5; DB 21; Length 940;
 Best Local Similarity 25.3%; Pred. No. 2;
 Matches 44; Conservative 30; Mismatches 59; Indels 41; Gaps 10;

QY 1 MGQHPAKSNDVRRIRIEGEL-----LNQLAGRMIPKGLTWSGKFPPTLDHVDHVQTM 54

Db 431 mqrh--rsnqlfngbhgmptqsfqgemgkfmksqgl-----qlyhngsqg 478

QY 55 EINTLQNGAMPAGARRVGLSNPTTQEIPOQWTPPE-DOKAREAFRRYQEERP---PE 110

Db 479 llsqlqggs-----kdmpprfskkgqlnadeislrpagsflmknqvpklqpq 526

QY 111 TTITPPSS-PPQKLPQ-GDDPLGNQSLLETH-PLYQSEPAVPVIRKPPKKK 161

Db 527 ltmippsqpprtqtpplggtqplg-----lkcnppllqekpaktakpppsakee 576

RESULT 9

AAW95400

XX ID AAW95400 standard; Protein; 407 AA.

AC AAW95400;

XX 26-MAR-1999 (first entry)

DE M. tuberculosis RP-factor-like gene product (MtubMTV043).

XX RP-factor; resuscitate; latent cell; growth-stimulation; receptor;
 KW convertase; vaccine; immunotherapy; prophylaxis; excipient; antibiotic;
 KW antimicrobial; tuberculosis; food; pharmaceutical; culture additive.

OS Mycobacterium tuberculosis.

XX W09855624-A1.

PN 10-DEC-1998.

PD 03-JUN-1998; 98WO-GB01619.

XX 27-MAY-1998; 98GB-0011221.

PR 04-JUN-1997; 97GB-0011389.

XX (UYWA-) UNIV WALES.

XX Kaprelyants AS, Kell DB, Mukamolova GV, Young DI;

PI Young M;

XX WPI; 1999-070270/06.

XX

PT New bacterial resuscitation factors - useful for stimulating latent
 PT bacteria and growth, particularly for use as antimicrobials, also
 PT vaccines comprising bacteria with mutations in resuscitation factor
 XX genes.

PS Disclosure; Fig 1A, 1D; 76pp; English.

XX The invention relates to RP-factors which are substances that can
 CC resuscitate dormant, moribund or latent cells, possibly also having
 CC growth-stimulating activity. Host cells containing a vector comprising
 CC the RP-factor or its receptor encoding nucleic acid can be used for the
 CC recombinant production of the RP-factor. RP-factors and their receptors or
 CC convertases, antibodies (against the RP-factor. RP-factors and RP-factor receptors
 CC or convertases), antagonists and agonists, are useful in vaccines and for
 CC immunotherapy, diagnosis and prophylaxis. They are also useful as
 CC excipients, generally as antimicrobials and especially for infections
 CC associated with latency. They can be used as potentiators of antibiotics
 CC such as isoniazid, streptomycin etc., in treatment of tuberculosis.
 CC RP-factor can also be used: to determine the microbiological quality of
 CC foods, pharmaceuticals, medical products; as culture additive for
 CC bacteria; to stimulate growth and/or to resuscitate microorganisms and to
 CC produce libraries of biomolecules and microorganisms (which may then be
 CC screened for useful products). Probes complementary to the RP-factor
 CC nucleic acid are used to identify and clone other RP-factor genes. The
 CC inventions may also be used to prevent bacterial resuscitation. Breaking
 CC dormancy with RP-factor facilitates detection, culture and enumeration of
 CC many bacteria. The present sequence represents a predicted amino acid
 CC sequence of a RP-factor-like gene product from M. tuberculosis.

XX Sequence 407 AA;

Query Match 10.3%; Score 90; DB 20; Length 407;
 Best Local Similarity 27.5%; Pred. No. 1.5;
 Matches 44; Conservative 14; Mismatches 68; Indels 34; Gaps 10;

QY 11 VRRIE-GGEILLN-----QLAGRMIPKGLTWS-----GKFPPLDHDHVDHVQTM---EEINT 58

Db 47 varcesggnwsintgngylgqlftqs--twaahgggafapsaqslasreqgiavgerla 104

QY 59 LQNGAMPAGARRVGLSNPTTQEIPOQWTPPEEDQKAREAFRRYQERPPPTTIPSS 118

Db 105 tqgrgawpv-cgr--glsnatprevlpasaamdapldaaav-----ngepaplapppad 155

QY 119 P-PQWKLQPGDDPLGNQSLLETHPLXYQSEPAVPVIRKTPP 157

Db 156 pappvelaandlpa-----plgeplpaapadpapp 185

RESULT 10

AA52470

ID AAB52470 standard; protein; 407 AA.

XX AAB52470;

XX 23-FEB-2001 (first entry)

DE Mycobacterium tuberculosis secreted protein #35.

XX Mycobacterium tuberculosis secreted protein; MTSP; vaccine.

XX Mycobacterium tuberculosis.

OS W0200066143-A1.

PN 09-NOV-2000.

XX 04-MAY-2000; 2000WO-US12197.

XX 04-MAY-1999; 99US-0132479.

PR 04-MAY-1999; 99US-0132503.

XX (PUBL-) PUBLIC HEALTH RES INST NEW YORK.

[illegible]

CC

Wang J., Wang Z., Wehrman T., Xu C., Xie A.T., Yang Y., Zhang J.

Zhao QA, Zhou P, Goodrich K, Dimanac R;
WPI; 2001-442253/47.
N-PSDB: AAI57996

Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -
Example 3: SEQ ID NO 1985; 10078pp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AAM38642-AA42213) with neurotrophic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Sequence 396 AA;

Query Match	10.0%;	Score 87.5;	DB 22;	Length 396;
% of Local Similarity	30.6%;	Pred. No. 2.6;		
Mismatches 30;	Conservative 10;	Mismatches 33;	Indels 25;	Gaps 3;

```

67 AGAGR-----RVGLSNTPQIPQPOW--TPEEDOKAREAPRRQEE 106
|||||
18 agagrfcflgseatrhlpharnhcgidsesqjlwepdfnprpkaskalsldfkryvtd 77
|||||
107 RPPETTTI----PPSGPPQWKLGPDGDDPLLGNQSLLE 139
|||||
78 rrlaetlaqlvqlkpsrpphllllecnnqpqltgalie 115
|||||

```

T 14
626
AAM40626 standard; Protein; 407 AA.

AAM40626;

22-OCT-2001 (first entry)

Human polypeptide SEQ ID NO 5557.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

Homo sapiens.

WO20015312-A1.

26-JUL-2001.

26-DEC-2000; 2000WO-US34263.

21-JAN-2000; 2000US-0488725.

23 REA 2000; 2000US-0598042;
09-JUL-2000; 2000US-0598042.

.ZTC0700 S00007'0007 T000 ET

